Capstone Project Final Report

Decaobonizing the “Last Mile” of E-commerce in Shanghai

Author: Zherui Li, Master of Urban Planning 2015
Department of Urban and Regional Planning at UIUC
Client: China Academy of Urban Planning&Design
Date: 5/13/2015

Executive Summary

Due to the surprising growth in E-commerce market, most urban areas in China have witnessed a significantly increasing quantity of parcel shipments transiting through the regions. Increase in urban freight traffic obviously exacerbates congestion and environmental problems. There is a strong need for urban areas, particularly those compact cities with high-density population, to implement effective strategies which minimize negative impacts of E-commerce logistics. Furthermore, though strategies to reduce logistics cost have attracted great research interest in past decades, “Green Logistics” which focuses on reducing carbon emissions is one of the newest topics in China. Therefore, the objective of this project is to propose viable “Green” freight solutions jointly in terms of reducing pollutant emissions and congestion. Specifically, two types of “Green” parcel-delivery strategies are considered in this research, i.e. vehicle routing optimization and collection point’s optimization. The former includes a routing and sequencing optimization model which could be applied by courier services and the latter includes a location-allocation analysis which could be implemented by the local government.

With the 2014 commercial survey data obtained from ShunFeng Courier Company and the demographic data obtained from the China Academy of Urban Planning and Design (CAUPD), a logistic study is first performed to investigate the current parcel-delivery network in Shanghai. Generally, “door-to-door” delivery, which represents that couriers directly bring parcels to customers, is the dominant delivery mode in Shanghai. In addition, couriers select either “the shortest” or “the quickest” set of delivery routes based on personal experiences. “Collection Points” system (CP), which represents that customers go to certain CPs to pick up parcels by themselves, however, are underutilized. Two facts contribute to the underutilization of CPs: first, no empirical study justified that CPs’ benefits could outweigh the costs; second, neither courier services nor the local government is willing to bear the expense of relocating existing CPs or adding new CPs. Results from this study provide the basis for the modeling assumptions in the following vehicle routing problem (VRP) model and the location-allocation analysis:

- For VRP model, a static model is developed to find the optimal sequence of customers being visited and the optimal routes to finish all delivery tasks, by quantifying the effects of road infrastructure, traffic condition, and weather on carbon emission amount per mile of delivery vans. Furthermore, a dynamic model is proposed to optimize departure time at each customer point, by considering the real-time traffic condition. The latter model uses the optimization result of the former model as inputs.

- For location-allocation analysis, a comparison between “door-to-door” delivery and “Collection Points” system is made to justify the advantages of CPs in reducing carbon emissions. In addition, an analytical framework for evaluating the efficiency of CPs and relocating low-efficient CPs is presented.

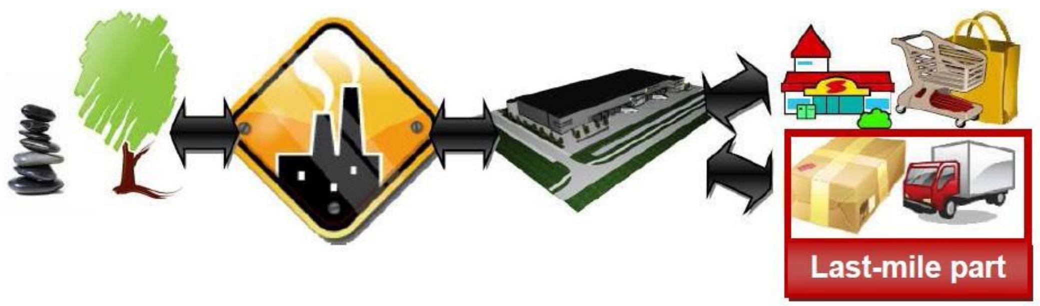
In sum, this project 1) provided an initial and good understanding of the current parcel-delivery system of Shanghai; 2) developed a framework for quantifying carbon emission cost and considering multiple factors influencing carbon emissions, including infrastructure, traffic, and weather; 3) presented a systematic way to periodically evaluate CPs and choose reasonable locations; 4) demonstrated that when all the conditions are the same, collection points system can reduce much more carbon emissions than the optimized vehicle delivery routes. The project is believed to help the local government promote Collection Points in Shanghai without adding financial burdens on courier services.

Table of Contents

Introduction to “Last Mile” delivery

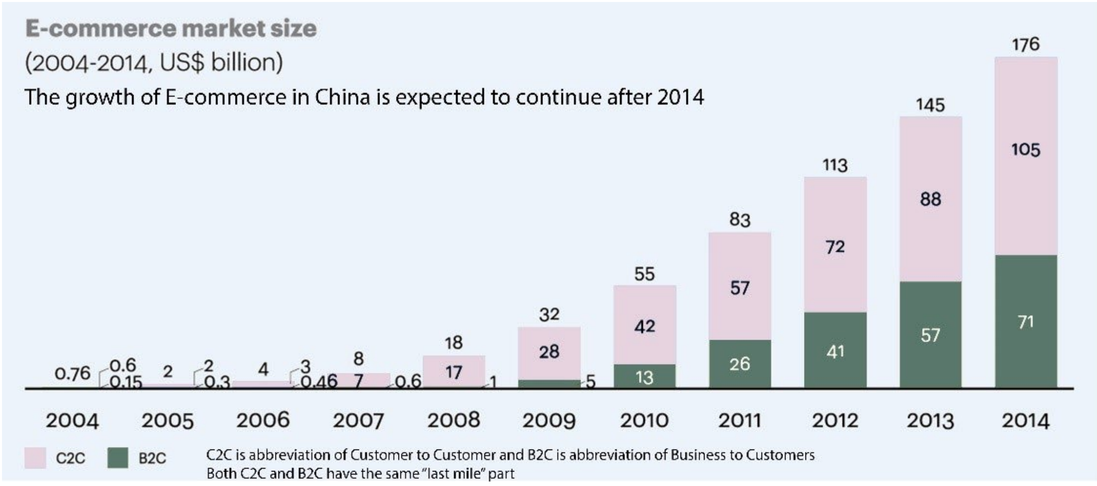
In China, since the last decade, the E-commerce market experienced by enormous growth. The E-commerce spending has risen from $0.76 billion in 2004 to more than $176 billion in 2014 and is estimated to continue growth.[1](see graph 1) This growth significantly increased the number of direct-to-customer deliveries and resulted in the emergence of a complete supply chain including distribution, package and transport etc. The last part of the supply chain, which refers to delivery to the final recipient, is called “the last mile”.[2](see graph 2)

Graph 1 Supply Chain of Logistics



Source: Own reproduction based on De Smedt, Gevaers (2009)

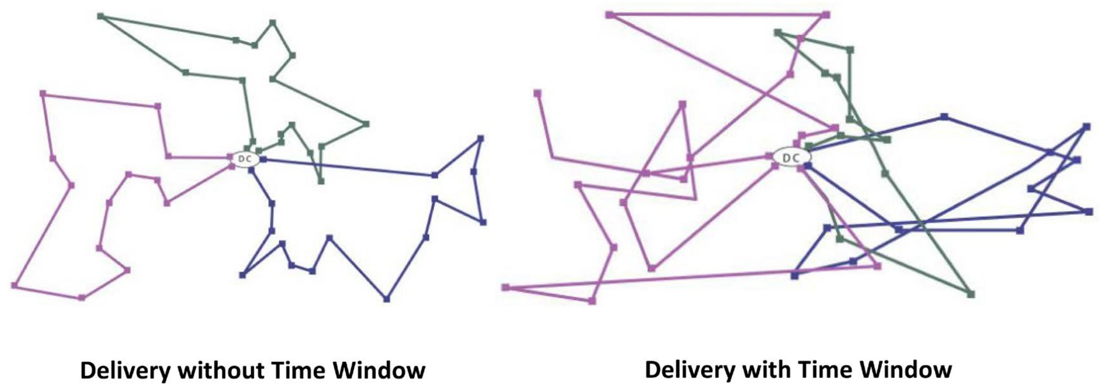
Graph 2 E-commerce Development in China



In general, two types of delivery systems are widely used in the “last mile” delivery: “e-commerce networked delivery”(END)system and “sustainable networked delivery”(SND) system. In the END system, otherwise known as door-to-door delivery, goods are first distributed from warehouse to a distribution center, and then conveyed by vans to each customer. If the good is not delivered to the customer at the first time, the courier needs several other attempts before returning the good to the warehouse. In the SND system, sometimes called self-pick-up system, products are first distributed from a warehouse to a distribution point, and then carried by vans to a collection point, where each customer come to pick up their goods by themselves.

The END system is considered as the most inefficient, expensive and polluting part of the whole delivery chain for two reasons. The first one is the high frequency of failed deliveries due to the “not-at-home” phenomenon. This phenomenon occurs when no specific time windows was set between couriers and customers. If customers are not at home, couriers have to return again. To avoid such phenomenon, most express companies require time window being agreed before delivery.[3] Time window(TW) refers a period of time within which the delivery service should be completed. There are two types of time window: soft TW AND hard TW. Soft TW means the time window could be violated to acceptable degree and the violation will be converted to some monetized punishments on couriers. Hard TW means the time window must be followed and if the violation occurs, customers will directly refuse to receive parcels. However, either soft TW or hard TW results in a same problem.[3] Like graph 3 shows, the delivery with time window results in a so called “Ping-Pong effect”, which means a surprising increase in driven kilometers and subsequent CO_2 emissions.

Graph 3 Types of Time Window in Delivery



The second one is the “empty run” syndrome. Online shopping heat is seasonal, especially accompanied with festival sale. Thus, during ordinary weekdays, even in densely populated cities, delivery vans’ loading capacity cannot be fully utilized. Additionally, for sparsely populated areas, customer density may even not be able to operate at an acceptable level of costs. After all, no courier wants to drive 20km for delivering one parcel. In one word, the above two reasons make door-to-door delivery quite difficult to be efficient.

Compared with END system, SND is much more efficient and environment-friendly. Previous studies about delivery of online books justified the advantages of SND in reducing CO_2 emission. Kim and colleagues(2008, 2009) designed several collection points for picking up books by customers themselves in the United States. They compared the emissions of the proposed “sustainable networked delivery”(SND) system and “e-commerce networked delivery”(END) system in delivering 100 books to 100 customers. Their 2008 study indicated that the CO_2 emissions of END are 12 times those of the SND system and their 2009 study showed that difference was 7 times. Zhang(2013) conducted a similar study in China’s second tire cities, his study showed that though SND system in those cities has greater environmental impacts than that in UK and US, SND still generates less CO_2 than END. Though the latter study took into account CO_2 emissions produced by customers on the way to collection points while the former one did not, SND’s advantages should not be doubted since this system could ensure each parcel be successfully delivered by first time and the loading capacity of collection points be fully utilized.

Significance of the project

Three factors justify the urgency to optimize Shanghai's online goods shipment system in terms of reducing carbon emissions. First of all, China has a huge-size logistic industry. In 2014, the total spending on logistics in China was up to $1600 billion dollars, occupying 17% of the whole GDP, and this rate is twice that of US. Take Shanghai for example, Shanghai express companies have 2,949 distribution centers, 54,712 employers and 17,400 delivery vans. The volumes of delivery parcels grow at an annual rate of 40%. Assuming each delivery van drive 100km per day, the whole express system will generate about 350 ton CO_2, occupying 2% of the total CO_2 emission generated by automobiles per day. More importantly, Shanghai's city core area is experiencing serious traffic congestion every day. The average vehicle speed on express way during peak hours is 40km/h and V/C ratio of the section of high ways falling within city cores is up to 0.8. Though there are no explicit data extracting the impacts generated by express vans, most parcels were delivered to customers within city cores, which definitely exacerbated the traffic congestion.

In addition, most of the literature and models involving "last mile" of products delivery revolves around minimizing total transportation cost and nearly all of these models were developed by private express companies. The major objective of those models is either to minimize the total travel distance of a fleet vehicles or to minimize the total time spent on roads. These models play limited role in reducing carbon emissions and benefiting the society as a whole. In fact, most Chinese express companies even do not apply those transportation models and couriers choose delivery routes based on observation and experience, which makes massive delivery routes of different companies are overlapped and thus more congested. The final factor is that government's intervening with logistic is low efficient. During past three years, Chinese big cities built a lot of large-size logistics parks to locate express companies and uniformly distribute parcels. Government officials believed that the efficiency of delivering parcels could be improved significantly if the location of logistic parks was scientifically determined. However, after such logistics parks were built, there occupancy rate were low. The underlying reason is that express companies located their own distribution centers for proximity to customers while centralized logistics parks fully disobey market laws. Therefore, we meet a dilemma: can we mitigate the negative environment impacts derived from "last mile" delivery without hampering express companies' interests?

Objectives and methodology

This project will try to solve that dilemma through accomplishing two objectives: 1. For express companies, I will build one transportation model which could optimize delivery routes in terms of minimizing carbon emissions based on the location of each company's distribution centers and customers. This model consist of a static submodel as well as a dynamic submodel, and express companies could choose to use one of them or both of them based on their spending on environment protection. Express companies can print out delivery routes and distribute them to couriers before each day's delivery. 2. For government, I compare the total volume of carbon emissions generated by END system and SND system. The result shows that even though the optimized delivery routes would be applied in END system, SND system still produces less carbon emissions. In addition, I also design a GIS analysis approach for governments to choose best locations for collection points based on historic customers' data. I believe that the comparison between END and SND as well as the GIS approach could encourage the government to issue policies popularizing SND system.

This project is divided into two sections: "Green" vehicle routing optimization and Collection points' location optimization. The research methodology used in the first section includes mathematic model building, algorithm design and simulation. Initial codes are written for Matlab to apply the proposed models in simulation and GIS is used to present the simulation results. In the second section, research methodology includes spatial analysis, allocation-location analysis and case study. GIS extension tool box is used to conduct all analysis and present findings. Besides, literature about strategies of building efficient collection points are reviewed to supplement findings. The data and materials used in this project are listed as follows:

- Land Use map and population distribution by community of study area
- Address location of existing distribution centers and collection points that fall in study area
- Street network shapefile of study area
- Historic Traffic Data by zones of study area
- Randomly generated customers based on land use type

These data have several limitations: first, I could not acquire information about more detailed population distribution within land tracts. Thus, population is assumed to evenly distribute within each community. Second, I could not get traffic data on each roads of study area and the smallest unit that can collect data is traffic zones. Therefore, streets that fall in the same traffic zones are assumed to have the same traffic condition during certain period of time. Third, due to business privacy, authentically historic customers' information including customers' address location, demand and time window could not be acquired. But the above limitations have been avoided to the largest extent through making reasonable assumptions, and the detailed research steps will be clarified in later sections. The deliverables of this project are as follows:

- Optimized END vehicle routing map with static traffic condition as inputs. This map includes the sequence of each customer being visited and the routes for courier to reach each customer.
- Optimized END vehicle routing map with dynamic traffic condition as inputs. This map includes all information mentioned in last map and the departure time as well as waiting time, if necessary, at each customer.
- Optimized SND vehicle routing map with dynamic traffic condition as inputs. This map allocates each customer to nearest collection points, and this map includes the sequence of each collection point being visited and the routes for couriers to reach it.
- A table that compares END and SND. This table will list travel distance, time spending and carbon emissions of deliverable no.2 and no.3 respectively.
- Service area map of the existing collection points. This map shows service area of each collection point based on 5-minute walking and 10-minute walking buffer. This map also ranks each collection points based on the number of customers that fall within their service area.
- Location map of potentially new collection points. This map shows that to maximize number of customers by one year within 10-minute walking service area of collection points, how many existing collection points should be relocated and how many new collection points should be added.

Introduction of Study area

The study area of this project is Huangpu District of Shanghai City. In fact, the online consumer market is currently concentrated in regions, with the top 10 cities all located on the east coast. Shanghai is the largest market with $2.6 billion in total online spending from 2009 to 2010, accounting for 8.7 percent of total spending. (see graph 4) In 2012, the total number of parcels generated in Shanghai was 599 million, equaling to 25 parcels/person/year, while the counterpart of the whole China was 10 parcels/person/year. As the cultural, social and commercial core of Shanghai, Huangpu district consists of 10 communities and its total population is 678,670 by 2013. With the area of 12km2, its population density is up to 56,555 person/km2, becoming the most populated district of Shanghai. In addition, the density of street network in Huangpu District gets close to that of the core area in Paris and it is one of the most congested districts in Shanghai. Thus, Huangpu District is an ideal place to conduct this project and the results of this project will play a significant role in mitigating traffic congestion and reducing carbon emissions within this district. There are 13 express companies that have business in Huangpu District, and the basic data used in project was obtained from ShunFeng Express which is the largest express company in China.

Graph 4 Top 10 Cities in terms of Online Spending

Like the whole Shanghai city, the current "last mile" delivery system in Huangpu District is "door-to-door"delivery(END)-based and supplemented by collection points system(SND). Specifically, couriers departure from the depot, following a soft time window which was negotiated with customers, visit a series of customers and then come back to the depot after finishing all delivery tasks. In China, time window of delivering goods is obscure and complicated, and thus following characteristics have to be clarified first in order to better simulate the delivery procedure later: customers do not have the right to set individual time window, but a soft time window still exists. For instance, when I purchased goods online, I could not require what specific time periods I expect to receive them. Instead, to maintain profits and reputation, most express companies make "24-hours" arrival as a tacit timing requirement. That means if I place my orders at 1:00 pm today, the express company will finish delivery before 1:00 pm tomorrow. However, since the courier will not tell me his specific arrival time and I may receive my parcels any time before 1:00 tomorrow, I had to be cautious when I typed in the delivery address. The reason is simple: unlike in US, Chinese couriers cannot drop off parcels at the boor when people are not at home due to high theft rate. In addition, Chinese couriers' bad service attitude forced customers to successfully receive parcels in first-time delivery as far as possible. Therefore, in the above example, since the parcel is estimated to arrive in daytime, I would make my workplace address as the delivery address. In fact, nearly half of the total parcels per year in Shanghai were delivered to commercial area while another half

fell in residential area. With this "smart strategy" being adopted by customers, the "not-at-home" ratio in China is actually quite low(about 3 of 80 cases), so all parcels are assumed to be successfully delivered within the first attempt. Details about the simulation's operation will be clarified in later chapters.

In Shanghai, there are formal and informal collection points, which play different roles in delivery system. Formal collection points were built by Express companies themselves and there are employees to manage each point. These collection points often locate near university campus, CBD and residential areas filled with young people, all of which are regarded as stable sources of parcel demand. Customers within these area walk a short distance to pick up goods at collection points. Compared with "door-to-door" delivery system, these collection points save much more labor and time cost. Due to the limited spending on maintaining collection points and lack of evaluating method, express companies are very cautious in increasing or relocating existing collection points. On the contrary, informal collection points are products of "not-at-home" syndrome. These points are usually grocery stores that locate within residential areas with good neighborhood relationship. Specifically, owners of these stores are familiar with parcel recipient and they are willing to receive parcels when the recipient are not at home. This not only helps couriers avoid another delivery attempt but also generates positive impacts on the stores' selling. Since the study area has a lot of neighborhood units that were built in 1980s, the number of informal collection points is huge and their distribution is compact. Integrating the informal collection points which are close to each other will probably improve the efficiency of collecting parcels, but this means significant increase in maintenance cost to expand storage space. In fact, if informal collection points could build a cooperative relationship with express companies in the following way: informal collection points provide storage space and express companies provide subsidize, both of them would be better off.

"Green" Vehicle Routing Problem Optimization for END system

"Green" vehicle routing problem(GVRP) evolved from ordinary VRP and their underlying optimization logics are similar. Two key questions involved in VRP optimization are which routes should be selected(also known as routing part)and in what sequence should these routes be traveled(also known as sequencing part). The purpose of ordinary VRP is to minimize the sum of different kinds of costs during delivery. The typically VRP model is listed as follows:

$$\mathrm{MinF} = \sum_{k \in K} BZ_k + \sum_{i \in N} \sum_{j \in N} \sum_{k \in K} Cd_{ij} X_{ijk} + \sum_{i \in N} PT_i$$

This model assumes that there are N customers and K delivery vans. Each delivery van departures from the depot to provide services for a set of customers through traveling a set of arcs. After finishing services, each delivery van should go back to the depot. B is the set-up cost(fixed charge) of each vehicle. C is the traveling cost per mile, which takes into account the amount of gasoline consumed per mile. is the distance of the arc travelled by the delivery van between customer i and customer j. P is the monetized punishment per minute for time window violation. is the length of time violation. Both and are both binary variable. means if the vehicles is used, then . Similarly, refers to whether the arc between i and j has been travelled. The goal of this model is to minimize the total cost F.

The method of building a GVRP model which considers minimizing carbon emissions as the purpose in previous literature is quite simple: transforming the above real distance into "virtual distance". These models assume carbon emissions are positively associated with gasoline consumption, and as long as gasoline consumption is minimized, the goal could be accomplished. For instance, one previous study points out that the amount of gasoline consumed per 3 minutes during traffic congestion equals to that of smooth driving for 1 km. Then, the real distance between customer i and j is converted to "virtual distance" by using following equation , where t means the congestion time. In addition, the rest variables like C and P etc. in the ordinary VRP model remain the same. This so-called GVRP model, in essence, is still a model of cost minimization and it has one obvious drawback. The drawback is that this model assumes that with the vehicle speed increase, the amount of carbon emission would infinitely keep decreasing.

However, the real relationship between vehicle speed and carbon emissions rate is illustrated by the following graph. The optimal vehicle speed with the lowest carbon emission rate is around 50km/h. When the speed decreases to approach 0, the carbon emission rate increases significantly. When the speed is over 50km/h and keeps increasing, the carbon emission will slightly climb. Therefore, to minimize carbon emissions, a model which uses vehicle speed as an independent variables could be much more reasonable. In this project, I propose two models using static and dynamic vehicle speed as inputs respectively. In addition, to better simulate the real delivery situation in Shanghai, only the existing depots will be taken into account, which means no new depots will be proposed.

- **Static "Green" Vehicle Routing Problem Optimization model**

The static GVRP model uses static traffic information as inputs and optimizes delivery routes to minimize the total amount of carbon emissions. Before the detailed model building process is clarified, we should notice that there are four factors affecting the volume of carbon emissions during delivery, which are delivery efficiency, loading ratio, traffic and infrastructure. The former three, which refer to "not-at-home", "empty run" and optimal travel speed, have been mentioned in previous sections. Since the "not-at-home" rate is quite low in Shanghai, the simulation will not take into account multiple delivery attempts. In terms of loading ratio, optimizing the loading ratio of each delivery van is part of the whole optimization model. Infrastructure refers to the hierarchy of roads. Typically, high-level roads has higher speed limit, better pavement and subsequent higher average travel speed. According to graph, an extremely high travel speed enjoys lower carbon emission rate than an extremely low travel speed. Thus, higher-level roads have more advantages in reducing carbon emissions.

The whole optimization includes three steps: the first step is to divide the whole study area-Huangpu District into several delivery districts. Each district is served by one depot and all such districts should not overlap. The second step is to further divide each delivery district into smaller delivery units and all delivery tasks within one delivery unit should be finished by one delivery van. The third step is to use a mathematic model to optimize delivery routes of deliver vans departing from the same depot to serve customers within the same delivery district. The former two steps are completed based on cluster analysis which integrates customers within particular search tolerance. The inputs are the Euclidian Distance(planar distance on the map) between each customer and the distribution center, and the Euclidian Distance between any pair of customers. The purpose of the second step is to assign each customer to the nearest depot, and the purpose of the third step is to cluster customers with similar characteristics. Both of these two steps can roughly reduce the unnecessary travel distance derived from overlapping

delivery districts and sparsely distributed customers. The last step will produce an accurate model which takes into account carbon emission rate and time window punishment.

Step 1. Production of delivery district

Since there are 6 existing depots for Shungfeng Express Company in Huangpu District, this step can be simply described as: assign m randomly generated customers to 6 depots. The method is listed as follows:

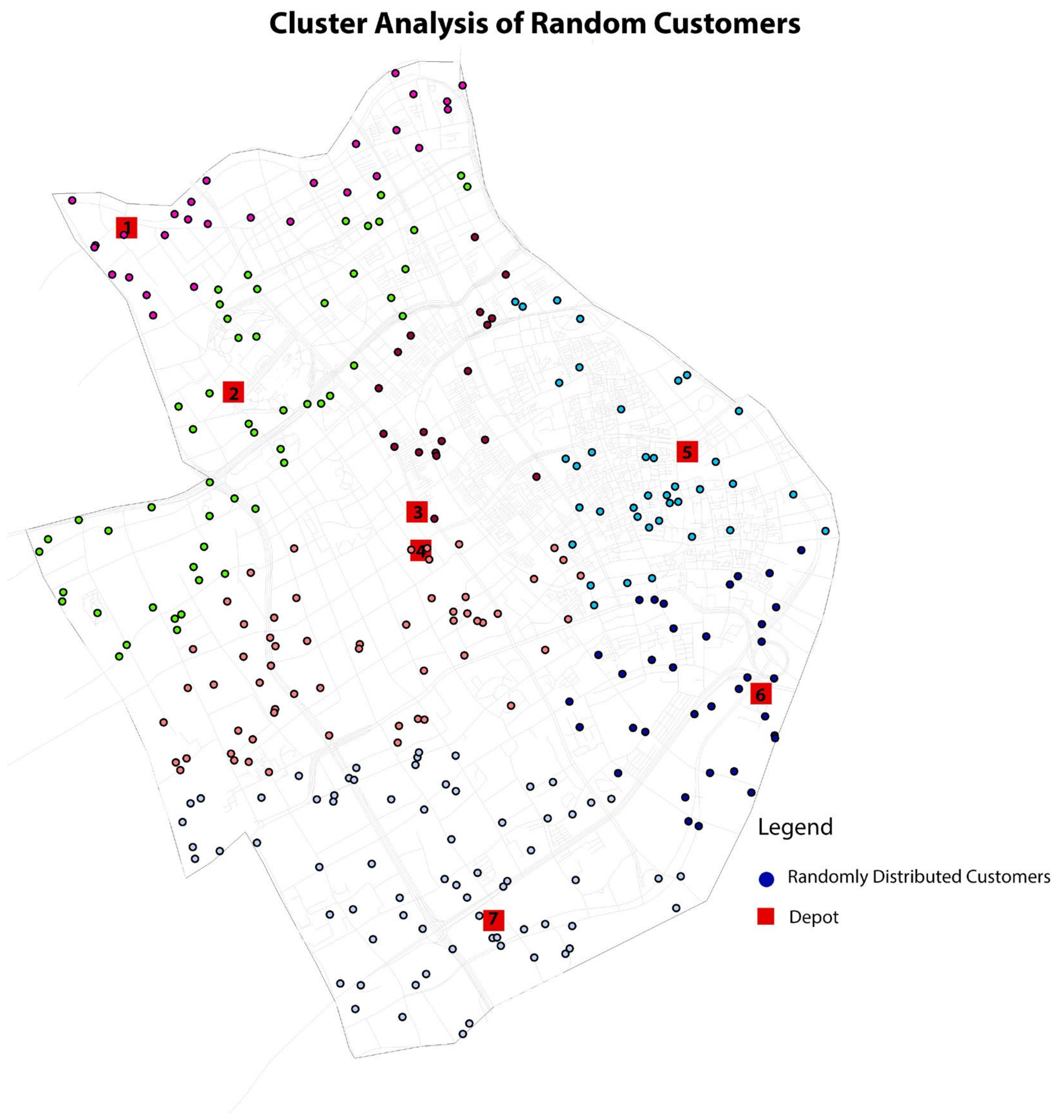
First, calculate the distance between m customers and the 6 depots

$$D = (X_i - Y_i)^2 + (X_j - Y_j)^2; 1 \le i \le m; 1 \le j \le 6$$

(Xi, Yi) is the coordinates of customers while (Xj, Yj) is the coordinates of depots, D is the Euclidian Distance.

Second, sort the customers from close to far, based on their distance to each depot. Assuming there are 300 hundred randomly generated customers in the study area and all of them need to be assigned to suitable depots. The result is illustrated by graph5 and different colors of customers define different delivery districts, in total, there are 6 delivery districts.

Graph 5 Production of Delivery Districts



Step 2. Production of smaller delivery units

This step integrates customers that fall in the same delivery district based on their “closeness degree”. “Closeness degree” is determined by the similarity degree of customers’ time window, geographic location and demand. It can be calculated by using following equation:

$$C = \left(\sum_{K=1}^{2} \left| X_{ik} - X_{jk} \right|^{2} \right)^{1/2}$$

X_i denotes customer i (i= 1, 2…m, i), X_{ik} denotes customer i’s Kth characteristics (k=1,2,3).
Since the unit of time window, geographic location and demand is different, I use mean standardization method to eliminate the unconformity among different units.

$$X_{ik} = \frac{X_{ik}{}'}{\overline{X_k}}$$

Besides, the number of delivery units is produced through dividing the total demand of all customers by the capacity of one delivery van.

$$K = \left[\frac{\sum_{i=1}^{m} d_i}{V} \right] + 1$$

K denotes the number of delivery units, $\sum_{i=1}^{m} d_i$ denotes the sum of all customers’ demand, [ ] is the rounding function
This could ensure the loading ratio of each delivery van can reach a relatively high level. Specifically, assuming that we meet an extreme case which is shown as follows:

$$\sum_{i=1}^{m} d_i = n \times V + \Delta$$

Δ *denotes a tiny number and n denotes the number of delivery vans*

The average loading ratio of each delivery van= $\lim_{\Delta \to 0} \frac{\sum_{i=1}^{m} d_i / n+1}{V} = \lim_{\Delta \to 0} \frac{(n \times V + \Delta) / n+1}{V} = \frac{n}{n+1}$

Even assuming n=2, the average loading ratio is up to 66.7%. When n increases, the ratio will also increase.

The whole procedure of producing delivery units is summarized as: first, determine K. Second, mean standardize time window, demand and geographic location. Third, run the K-means cluster analysis tool in software R. Fourth, double check the result of cluster analysis. Assuming there are three delivery vans in depot 5 and the capacity of each delivery van is 80. The information of customers are given and the delivery district of depot 5 need to be divided in to 3 delivery units. The result of dividing delivery units is illustrated by table 1:

Table1 Production of Delivery Units

Customer	Coordinate		Time Window		Demand	Cluster	Mean: 121.4919	31.2218	513.1579	904.7368
	X	Y	Start Time	End Time			Std: 0.0042	0.004715	82.17007	143.8794
0	121.4895	31.2159	480	840	10	3	-0.58763	-1.26658	-0.40353	-0.44994
1	121.4909	31.2199	480	1020	5	1	-0.25126	-0.41853	-0.40353	0.80111
2	121.4866	31.2320	720	900	5	2	-1.28157	2.161064	2.517244	-0.03292
3	121.4938	31.2197	480	1020	9	1	0.447205	-0.4591	-0.40353	0.80111
4	121.4960	31.2264	480	1020	3	1	0.986109	0.975466	-0.40353	0.80111
5	121.4890	31.2234	600	840	3	2	-0.70571	0.321959	1.056858	-0.44994
6	121.4938	31.2206	600	720	8	3	0.44521	-0.27491	1.056858	-1.28397
7	121.5012	31.2188	480	960	9	1	2.215882	-0.65637	-0.40353	0.384094
8	121.4871	31.2151	480	1080	7	1	-1.16245	-1.4391	-0.40353	1.218125
9	121.4911	31.2159	480	720	7	3	-0.19813	-1.27154	-0.40353	-1.28397
10	121.4873	31.2233	480	900	8	2	-1.1249	0.309945	-0.40353	-0.03292
11	121.4879	31.2275	480	1080	4	2	-0.97419	1.198855	-0.40353	1.218125
12	121.4894	31.2281	480	960	2	2	-0.61995	1.324764	-0.40353	0.384094
13	121.4901	31.2307	480	660	4	2	-0.45318	1.880417	-0.40353	-1.70099
14	121.4954	31.2262	900	1080	3	2	0.836419	0.933358	4.707823	1.218125
15	121.4967	31.2214	480	780	3	3	1.143802	-0.09084	-0.40353	-0.86695
16	121.4866	31.2186	480	1020	4	1	-1.29624	-0.68963	-0.40353	0.80111
17	121.4933	31.2197	480	840	1	3	0.314646	-0.45259	-0.40353	-0.44994
18	121.4888	31.2201	540	900	8	1	-0.76926	-0.37894	0.326665	-0.03292
19	121.4862	31.2324	480	840	7	2	-1.38137	2.235358	-0.40353	-0.44994
20	121.4910	31.2193	540	1020	9	1	-0.22002	-0.53168	0.326665	0.80111
21	121.4942	31.2176	480	780	9	3	0.54266	-0.89382	-0.40353	-0.86695
22	121.4953	31.2201	480	780	5	3	0.814185	-0.36649	-0.40353	-0.86695
23	121.4875	31.2205	480	600	6	3	-1.07315	-0.28271	-0.40353	-2.118
24	121.4928	31.2223	480	960	9	1	0.217208	0.102543	-0.40353	0.384094
25	121.4914	31.2254	480	1020	4	1	-0.12095	0.743465	-0.40353	0.80111
26	121.5027	31.2164	480	1080	9	1	2.585871	-1.15289	-0.40353	1.218125
27	121.4889	31.2320	480	960	10	2	-0.73975	2.148366	-0.40353	0.384094
28	121.4988	31.2239	480	1020	6	1	1.65765	0.434068	-0.40353	0.80111
29	121.4923	31.2189	480	720	1	3	0.094469	-0.63003	-0.40353	-1.28397
30	121.4975	31.2200	480	1020	7	1	1.333686	-0.38578	-0.40353	0.80111
31	121.4967	31.2176	540	960	1	1	1.1424	-0.91114	0.326665	0.384094
32	121.4932	31.2202	540	660	6	3	0.290833	-0.35649	0.326665	-1.70099
33	121.4879	31.2228	480	600	10	3	-0.98086	0.201252	-0.40353	-2.118
34	121.4920	31.2204	600	1020	9	1	0.009994	-0.31123	1.056858	0.80111
35	121.4924	31.2225	480	960	9	1	0.098573	0.135548	-0.40353	0.384094
36	121.4872	31.2162	480	1080	8	1	-1.15396	-1.20399	-0.40353	1.218125
37	121.4916	31.2186	480	960	5	1	-0.08227	-0.68203	-0.40353	0.384094

Step 3. Mathematic Model Building

This step is to accurately optimize delivery routes for delivery vans within one distribute district, and the method can be copied for optimization in other districts. The problem is summarized as: the three delivery vans of depot 6 need to provide services for customers assigned to them(the customer list conforms with that produced in step 2) and satisfy each customer’s time window. In order to minimize carbon emissions generated by each of the three delivery vans, the drivers need a map which shows the sequence of visiting customers and the route to reach each customer before departing from depot 6.

According to the content of previous paragraphs, we know that the carbon emissions rate has a complicated relationship with vehicle speed. In fact, the research result from UK Transportation Bureau demonstrates that the relationship between carbon emission rate and vehicles speed can be illustrated by following equation:

$$P_{ij} = \left(A_0 + A_1 * V_{ij} + A_2 * \left(V_{ij}\right)^3 + A_3 * \frac{1}{(V_{ij})^2} \right) \times L_{ij}$$

P_{ij}denotes the total amount of carbon emissions generated by delivery vehicles travelling between customer i and customer j, A0 A1 A2 A3 are constants 1 576; -17.6; 0.00117; 36.067respectively. V_{ij} is the vehicle speed between customer i and customer j, $L_{ij}$is the distance of street network between customer i and customer j.

However, in real delivery, vehicles cannot remain driving at a same speed between two customers, which means Vij keeps varying. To simplify the calculation of Pij , this model proposes the concept "virtual distance". Virtual distance quantifies the impacts of streets infrastructure and traffic condition on travel speed by following equation:

$$L'_{ij} = L_{ij} \times (1 + W_{ij} + R_{ij})$$

L'_{ij} *denotes virtual distance*, W_{ij} *denotes infrastructure and* R_{ij} *denotes traffic condition*

In other words, I assume that all delivery vans can travel between any pair of two customers at the optimal vehicle speed(50km/h) with the lowest carbon emission rate. But the really independent variables are infrastructure conditions and traffic conditions between any pair of customers. Thus, the amount of carbon emissions is illustrated by the modified equation

$$P_{ij} = \left(A_0 + A_1 * V_{ij} + A_2 * \left(V_{ij}\right)^3 + A_3 * \frac{1}{(V_{ij})^2} \right) \times (1 + W_{ij} + R_{ij})$$

The value of $W_{ij}$and$R_{ij}$is listed by following table

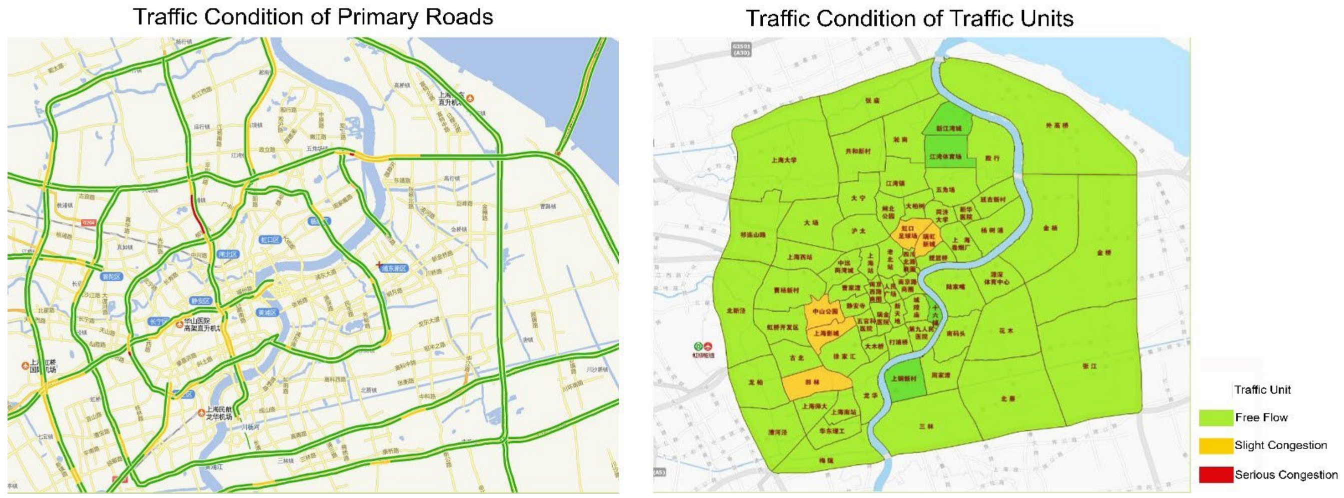
Table2 Quantified weights of Multiple Factors

Factor	Infrastructure Condition Wij			
	Primary	Secondary	Tertiary	Residential
Value Range	0	0.2	0.6	0.8
Factor	Traffic Condition Rij			
	Free Flow	Slight Congestion	Congestion	Serious Congestion
Value Range	0	0.3	0.6	1

For W_{ij}, streets' hierarchy equals to or higher than primary should apply the value 0 while streets' hierarchy equals to or lower than residential should apply the value 1. The information of streets' hierarchy is included in the street network shapefile.

For R_{ij}, Primary and higher-level roads can match their values with their traffic condition by checking the real time traffic information which could be accessed from the Shanghai Transportation&Travel Website. Streets whose hierarchy is lower than primary can match their values with traffic conditions within each traffic unit. Those low-level streets within one traffic unit are assumed to have the same value. Since this model only takes into account the static traffic condition, we can use historic information as inputs.

Graph 6 Traffic Information in Shanghai



The final model is listed below:

Explanation of terms:
$V_0$is the ideal speed of van(50km/h), which is a fixed constant.
K is the number of delivery vans
$L_{ijk}$is the real distance between customer i and j
$P_{ijk}$is the total amount of carbon emission generated by vans travelling between customer i and customer i
$L'_{ijk}$is the virtual distance between i and j
M is the capital cost of carbon emission per unit. The value of M is determined by Shanghai's carbon tax policy, which equals to $5 per ton
$Z_i$is the penalty cost for time window violation per minute
$ET_i$is the beginning of the time window of customer i
$LT_i$is the end of the time window required of customer i
$T_i$is the time when the van stops by the customer i
$T_o$is the starting time of the whole delivery service
T_e is the ending time of the whole delivery service

The final optimization model:

$$\mathrm{MinF} = \sum_i \sum_j \sum_k \left[\overbrace{A_0 + A_1 * V_0 + A_2 * (V_0)^3 + A_3 * \left(\frac{1}{V_0}\right)^2}^{P_{ijk}} \right] \times \overbrace{L_{ijk} * [1 + W_{ijk} + R_{ijk}]}^{L'_{ijk}} \times X_{ijk} \times M$$ **Formula 1.1**

$$+ \sum_i Z_i * \{\max[ET_i - T_i, 0] + \max[T_i - LT_i, 0]\}$$ **Formula 1.2**

S.T.

$$\sum_i \sum_j \sum_k X_{ijk} = K$$ **1.3**

$$T_{ijk} = \frac{D_{ijk} * [1 + W_{ijk} + R_{ijk}]}{V_0}$$ **1.4**

$$T_j = T_i + T_{ijk}$$ **1.5**

$$T_o + \sum_i \sum_j \sum_k T_{ijk} \le T_e$$ **1.6**

$$X_{ijk} = \begin{cases} 1, & \text{if the van drives from customer } i \text{ to customer } j \text{ by using the route } k \\ 0, & \text{Otherwise} \end{cases}$$ **1.7**

1.1 means the cost derived from the total carbon emission generated by goods delivery
1.2 means the penalty cost derived from providing service too early or too late
1.3 means there are N vans providing service in total
1.4 means the time that the van spends between customer i and j
1.5 means the time when the van arrives customer j
1.6 means all delivery has to be done before the delivery center closes
1.7 means the van has to derive on the k route between i and j

Step 4. Genetic Algorithm

The model above actually has infinite answers since there would be infinitely potential roads between customer i and j. Based on empirical research, there are several ways to solve this problem. A typical one which is used for solving the vehicle routing problem with time windows is the Genetic Algorithm. Following is a simple example to explain how Generic Algorithm works:

Task: To calculate the maximum number of f(x1,x2)=x12 + x22
X1 ={1,2,3,4,5,6,7}
X2 ={1,2,3,4,5,6,7}

Step 1: Coding
Since the Matlab can only tacking codes rather than numbers, we have to find an appropriate coding way to represent the numbers. For example, by using the 3-digits binary coding, one potential answer for the equation is y=101110, which equals to y=[5,6]

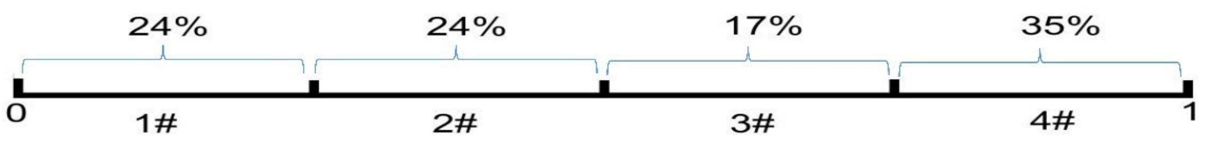
Step 2: Generate the first-generation answer
We need to randomly generate several potential answers as the first-generation which could be taken into later steps' calculations. For example, we choose four answers 011101, 101011, 011100, 111001 as the initial generation.

Step 3: Calculate the fitness
Fitness means the possibility of one answer to be remained in the next generation. For example, in the above task, our goal is to maximize the value f, thus, the bigger data has more possibilities to be chosen to participate into future calculations. On the contrary, smaller data would be abandoned.

Step 4: Choose
This step sets a rule which can guide us to choose "survival answers". For example, in this case, we first need to sum up the fitness of all the four initial answers and denote it as $\sum y$. Then, we need to calculate the fitness of each answer and denote it as y_i. The possibility for each initial answer to survive is thus $y_i/\sum y$. Since the sum of all ratios is 1, we can randomly generate a number between 0 and 1 for four times. The random number falls in which ratio range, that corresponding answer would be chosen. Details are attached in table3:

Table3 Rationale of "Choose" step for Genetic Algorithm

First Generation	X1	X2	Fitness(y)	Ratio to be chosen	Times to be chosen	Result of being chosen
"011101"	3	5	34	0.24	1	"011101"
"101011"	5	3	34	0.24	1	"111001"
"011100"	3	4	25	0.17	0	"101011"
"111001"	7	1	50	0.35	2	"111001"



(The distribution of ratio is accumulative and sequential. For example, if the first random is 0.19, it matches the ratio region of the first answer, which is 011101.)

Step 5: Cross
This step uses the “result of being chosen” column to generate new answers. The way is to randomly determine the digits which could be exchanged. Details are attached in table 4:

Table4 Rationale of “Cross” step for Genetic Algorithm

1	011101	1-2 3-4	1-2: 2 3-4: 4	011001
2	111001			111101
3	101011			101001
4	111001			111011

We can see that in the new answers, both “111101” and “111011” have a higher fitness than the previous results.

Step 6: Mutate
The last step is mutation, which means we need to totally change numbers on certain digits to generate more radical generations. Specifically, in this case, since it uses binary coding, we could determine several 0 to become 1 while some 1 turns into 0. Details are attached in table 5 and 6:

Table5 Rationale of “Mutate” step for Genetic Algorithm

Results after cross	Mutation	Results after mutate	New Answers
"011001"	4	"011101"	"011101"
"111101"	5	"111111"	"111111"
"101001"	2	"111001"	"111001"
"111011"	6	"111010"	"111010"

After this step, when we calculate the fitness again, we could find that the whole fitness goes up dramatically. Actually, if we like, this new generation could be regarded as the “first generation” and to repeat the steps 3-6. Finally, we will get a generation with the most desirable answer.

Table6 Rationale of “Mutate” step for Genetic Algorithm

	X1	X2	Fitness(y)	Ratio to be chosen
"011101"	3	5	34	0.14
"111111"	7	7	98	0.42
"111001"	7	1	50	0.21
"111010"	7	2	53	0.23
Total			235	1

Application of Genetic Algorithm in my project

Though the binary coding way is easy to be calculated, it is hard to be read by people. Thus, I chose using the natural numbers to code each customer. For example, I have 7 customers and 1 delivery center(depot). The delivery center is coded as “0” and customers are “1,2.3…..7”. The whole delivery process is “012345670”, which means our van starts from the delivery center and goes back to the delivery center after work. Since there are infinite number of available roads between each two customers, I narrow down this range to three roads with shortest distances. In other words, between each two customers there are road a, b, c for choosing. Thus, the final coding could be “0a1c2c….8b9c0”. Notice: the repeating “a” does not mean a same road, each pair of two roads have three potential roads coded as a, b, c to link them.

The fitness should be 1/F(formula 1.1 and 1.2) since the bigger fitness is, the more possible that delivery route is to “survive” in later calculations.

Choose

Following the steps described in the above example, I randomly chose 10 possible delivery routes. Then, I calculate each’s 1/Fi and the whole 1/F. By dividing them and randomly generating number between 0 and 1 for 10 times, I could finish this step.

Cross

In this step, I will exchange certain section between two routes. In other words, the 10 delivery routes generated from the last step will make 5 pairs and each of them will exchange sections.

For example, following is a specific process
1 6 3 4 5 2 7
2 1 7 6 5 4 3
After crossing, the new delivery route would be
1 6 3 6 5 4 7
2 1 7 4 5 2 3
Since we get repeating numbers, I change the one outside the crossing section to be a different number, so, the final delivery route would be
1 2 3 6 5 4 7
6 1 7 4 5 2 3

Notice: the corresponding roads linking each customer will also be exchanged, which means the customer point and the corresponding roads would act as a whole when crossing.

Mutate
I will randomly choose two customers to exchange their location within one delivery route. For example, the previous 1 2 3 6 5 4 7 would be 1 6 3 2 5 4 7, the mutation points of the 10 delivery routes would be totally different and random.

Simplification of the exploration process

Assuming there are N customers, and each customers have m available roads to get to them. In the Matlab, if I record each road's information in one row, there would be mN2 rows. For example, 0-1(m roads),0-2(m roads), 0-3(m roads)……0-N(m roads); 1-2(m roads),1-3(m roads)….1-N(m roads). To simplify the calculation, I assume the round trips between two customer points would have the same conditions. For example, if the start point is 1 and the destination point is 5, the delivery information would be the same from point 5 to point 1, which means I would eliminate the repeating information. By doing that, I will have M rows:

$$M = \sum_{i=0}^{N-1} (N * M - M * i)$$

And if I rank the new matrix by the ascending sequence, the row number will have following characteristics:

$$Row = \sum_{0}^{i-1} (N * m - m * i) + (j - i) * m + k - m$$

i means the start point and j means the destination point. K means the road chosen between i and j. row is the number of row which record the corresponding information. Notice: i=min(start point, destination point), j=max(start point, destination point). For example, if we wanna check the route information from point 4 to point 2, the start point i should be 2 while the destination is 4.

Step 5. Simulation

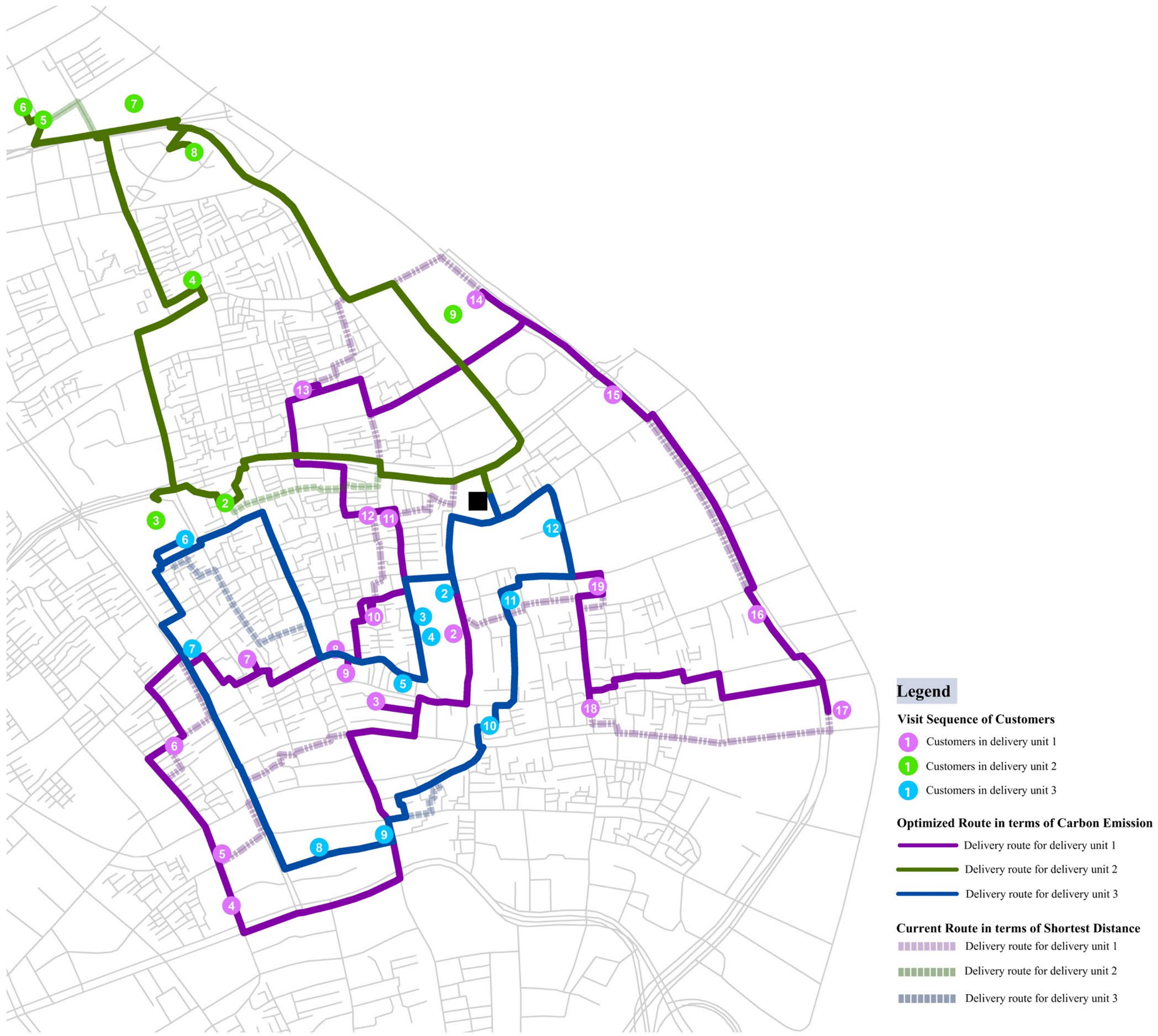
The simulation takes the result from step 2 "production of smaller delivery units" as inputs and applies the model and algorithm designed in step 3 and 4 respectively. Like the following map shows, three delivery vans departures from the depot 5 to serve 37 customers in total. The map shows the sequence of customers being visited for each delivery van and the routes that they should select. In addition, to justify the effectiveness of my proposed model, this map also shows the delivery routes selected by the current delivery optimization model which regards choosing shortest travel distance as optimization goal.

As table 7 illustrated, each delivery route produced by my model has less travel time and generates less carbon emissions than its counterpart of the current model. In fact, the three routes reduce 52 g carbon emissions, occupying 3% of the total amount. Since there are only 37 customers in the simulation, the optimization rate could be significantly large in real life. For example, the same-size area has about 28,000 parcels per day. If the optimization rate remains constant, the proposed model could reduce 39,000 g carbon emissions per day and 14.2 ton per year.

Table7 VRP Optimization Model Results

		Virtual Distance(meters)	Carbon Emissions(g)	Travel Time(minutes)	Travel Distance(meters)
Route 1	Current	12,930	802	15.5	7,815
	Optimized	12,479	774	14.9	8,246
Route 2	Current	7,412	460	8.9	4,532
	Optimized	7,143	443	8.6	4,608
Route 3	Current	7,651	474	9.2	4,421
	Optimized	7,526	467	9	4,549

Graph 7 Optimized Vehicle Delivery Routes in terms of minimizing Carbon Emissions



- **Dynamic "Green" Vehicle Routing Problem Optimization model**

Unlike the static "GVRP" part, this section describes and solves a green vehicle routing problem with time-dependency congestion. In fact, the static model that I proposed previously can be summarized as follows: assuming the distance and other initial road conditions of potential paths between each two customers are given, applying the model and the Genetic Algorithm which I designed could find a set of paths to minimize the carbon emission cost and time window violation cost. However, if there are massive delivery vans travelling in a city everyday, the static "GVRP" model which uses the initial and static congestion conditions as inputs would be unreasonable. A possible scenario is that an initially smooth road was selected by the model and Genetic Algorithm as a perfect route. Then, the route would experience an influx of delivery vans, turning into a congested road soon. When the delivery van drivers hold the route map distributed by managers before departure, they will find it is absolutely useless. In fact, the real world requires us to propose a dynamic model to flexibly respond to the real-time traffic flow.

The literature about time-dependency GVRP(TDGVRP) is few, most of which focuses on how to optimize the departure time at each customer point. For example, I had to deliver goods to two customers, one is in TBH and the other one is in Union. I first arrived TBH during peak hours but I did not provide service immediately since the road between TBH and Union was congested at that time. I chose to wait 5 minutes in TBH to avoid the congestion ahead. As mentioned in previous sections, since the carbon emission rate is associated with the vehicle speed, my "strategy" would somewhat reduce the carbon emission produced of my travel between TBH and Union. Though the most ideal method to deal with dynamic traffic is to change routes in time, but that requires a lot of cost on installing digital devices for each delivery van. In fact, a more feasible optimization way in reality is that each delivery van still have a planned delivery route before departure and the manager can tell couriers how long they should wait at each customer point to avoid congestion ahead based on the real-time traffic information.

Step 1. Mathematic Model Building
To make the model more practical and to utilize findings got from the previous parts, two assumptions are made:

1. Customers' information such as time windows, coordinates are not changed. Formulas that illustrates the relationship between carbon emission rate and vehicle speed are still valid.
2. Depot's information such as the number of delivery vans (3), each van's capacity (80) and the location of the depot (depot 5) is not changed.
3. For each delivery van, the sequence of customers being visited and the delivery route connecting each pair of two customers is given. Both of them are the same with the result of static GVRP model in last section.
4. To clarify the model-building process, only one delivery van and its corresponding customers will be included in the description of problem and model building part. The final optimization will include all the three vans.

Generally, the purpose of this model is to further optimize the result of the static GVRP model through figuring out the waiting time and subsequent departure time at each customer point.

Explanation of terms in the dynamic GVRP: The customer set is C= ($C_0, C_1 \ldots, C_N$), and C0 denotes the depot 5. Each vertex in C has a service time S_i>0, and a service time window[ET_I,LT_I]. In total, we have N customers. The arrival time of the van at customer i is$A_i$and the departure time isB_i. The selected route by the static GVRP model between each two customers has a distanceD_{ij}, and a travel time $T_{ij}(B_i)$, which is a function of the departure time from customer i. D_{ij}denotes the real street network distance rather than "virtual distance", "virtual distance" will not be used in this model. M is the cost of carbon emission per mile.

Description of the dynamic GVRP: The whole planning horizon equals to the working time[ET_0, LT_0] of the depot. The planning horizon is divided into M time periods and the length of each period is the same. Each period is defined by an interval T^m=[t^m,t_m]. Vehicle speed in one interval remains constant, but speed in different intervals may be different. For a vehicle that leaves at time Bi and travels between customer i and j, the vehicle would have a set of varying speed$V_{ij}(B_i)$={$V_{ij}^m(B_i)$,$V_{ij}^{m+1}(B_i)$,$V_{ij}^{m+2}(B_i)$………... $V_{ij}^{m+n}(B_i)$}. This set means that the vehicle's departure time $B_i$belongs to time interval $T^m$and thus the vehicle's departure speed isV_{ij}^m. Similarly, since the arrival time$A_j$belongs to time intervalT^{m+n}, the vehicle's arrival speed isV_{ij}^{m+n}. That is to say,$t^m \leq B_i \leq t_m$and $t^{m+n} \leq A_j \leq t_{m+n}$. In total, it needs n+1 time intervals with varying vehicle speed to cover the whole travel between i and j.

Except the speed, distance intervals is a function of departure time$B_i$as well. The distance traveled by the vehicle in each time interval is denoted the set$L_{ij}(B_i)$={$L_{ij}^m(B_i)$,$L_{ij}^{m+1}(B_i)$,$L_{ij}^{m+2}(B_i)$……. $L_{ij}^{m+n}(B_i)$}. The formula of the total carbon emission generated by vans traveling between customer i and customer

$$P_{ij} = \sum_{m}^{m+n} \{A_0 + A_1 * V_{ij}^m + A_2 * (V_{ij}^m)^3 + A_3 * \left(\frac{1}{V_{ij}^m}\right)^2\} * L_{ij}^m$$

The difference between the static model and the dynamic model in terms of applying this formula is that in the static model, $V_{ij}$is a constant(50km/h), and vehicles traveling at this speed could generate the least amount of carbon emissions per mile. The only independent variable in static model is the "virtual distance" which takes into account traffic condition and infrastructure condition. On the contrary, in the dynamic GVRP model, the speed $V_{ij}$and the distance $L_{ij}$are both variables and traffic congestion is an outcome rather than input.

In addition, although theoretically, the departure time$B_i$equals to the sum of arrival time $A_i$and the service timeS_i, we allow couriers to wait at customer i to avoid congestion ahead. Thus, the service start time may not necessarily be the same as the arrival time, which means$A_i + S_i \leq B_i$. Here, I denote the service start timeR_i, so$R_i \geq A_i$and$R_i + S_i = B_i$.

The final model of the dynamic GVRP

$$Minimize\ F = \sum_{(i,j)\in C} X_{ij} * M * P_{ij}(B_i)$$

note:$X_{ij}$is a binary variable indicating whether the vehicle travels from customer i to customer j($X_{ij} = 1$). Otherwise,$X_{ij} = 0$.

$$\sum_{i=0}^{n} X_{ij} = 1 \quad 1.1$$

$$\sum_{j=0}^{n} X_{ij} = 1 \quad 1.2$$

$$\mathrm{ET}_i * \sum_{(i,j)\in C} X_{ij} \le R_i \quad 1.4$$

$$\mathrm{LT}_0 * \sum_{(i,j)\in C} X_{ij} \ge R_i \quad 1.5$$

$$X_{ij} * (R_i + S_i + T_{ij}(R_i + S_i)) \le R_j \quad 1.6$$

Constraints 1.1 and 1.2 ensure that every customer would be visited only once. Constraint 1.3 ensures that the vehicle leaves from the depot only once. Constraints 1.4 and 1.5 mean the service start time must follow the time window requirement. Constraint 1.6 ensure there would be enough travelling time between customer i and j.

Property of the TDVRP model

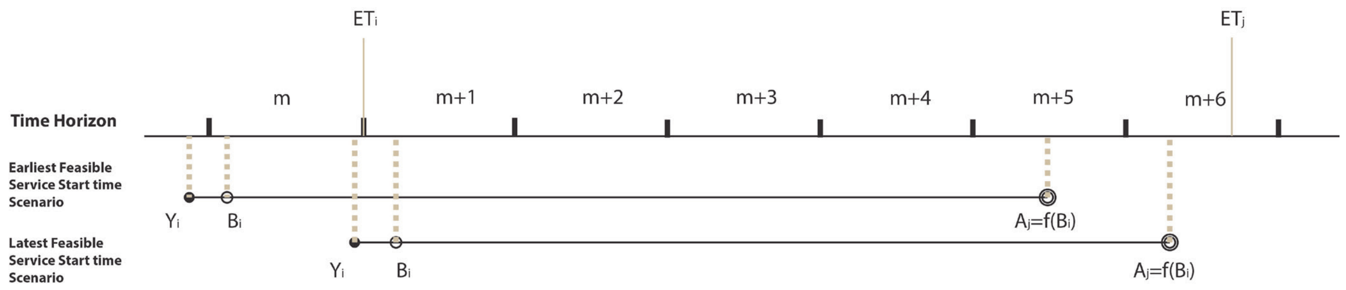
The purpose of the dynamic GVRP model is to optimize the departure time from each customer in a given route. But, the solution of the model seems complicated since the change of departure time B_i would affect a series of arrival time $A_j$later. In other words, since each customer has a time window, we cannot unlimitedly delay the service start time at one customer, which may make couriers violate time window of next customer. Thus, I denote $\hat{}R_i$ and $R_i\hat{}$ as the earliest and the latest feasible service start time at customer i respectively. The simplest way to solve the optimization model is to try all time points in the time interval[$\hat{}R_i$+S_i, $R_i\hat{}$+S_i]asB_i. However, since the time interval is a continuous domain which consists of infinite number of time points, it is unable to obtain the answer by using the simplest way.

In fact, the model solution can be simplified significantly when we notice the essence and the purpose of the optimization model: assign the longest travel distance to the time interval with relatively lowest carbon emission rate. In other words, if there was a time interval with quite low carbon emission rate, we need to ensure delivery van could "fully utilize" such period of time. Following is a straight example to justify this argument. Like the graph demonstrated, there is a pair of customers i and j. Vertex i has the earliest feasible service start time $\hat{}R_i$ and the latest feasible start time $R_i\hat{}$. Consequently, all possible scenarios about the leave time at customer i are defined by the union of $F(\hat{}R_i + S_i) \cup F(R_i\hat{}+S_i)$. That is to say, possible scenarios are covered by 7 time periods (m, m+1…m+6) with varying carbon emission rate.

Further analyzing the table which lists carbon emission rate for the above 7 time periods, we could find that there are two time periods(m+1 and m+5) with lowest carbon emission rate. Thus, without stops on the midway of the delivery route, lowest carbon emission scenarios could be found either for departures

time that coincide with the beginning of the "m+1" interval or for departures time that results in arrivals at the end of the "m+5" interval. The reason is that both of the above two scenarios "fully utilize" the "most desirable" time interval. (see graph 8) Generalizing this finding to the whole model, to obtain the optimization value, we just need to try time points which coincide with the beginning time of time intervals defined by [$\hat{}R_i$+S_i , $R_i\hat{}$+S_i] and points which result in arrival time coincide with the end time of time intervals defined by[$\hat{}A_j(\hat{}R_i+S_i)$, $A_j\hat{}(R_i+S_i)$]. This strategy will significantly reduce the number of our calculation.

Graph8 Successful Utilization of Time Intervals with Low Carbon Emission Rate



Time Interval	Carbon Emission Rate
m	400
m+1	200
m+2	400
m+3	1800
m+4	600
m+5	200
m+6	600

Step 2. Algorithm Design

Based on the property of the dynamic GVRP model, the first step I need to finish is to define$\hat{}R_i$ and$R_i\hat{}$. Since each customer is assumed to have a time window, $\hat{}R_i$just equals toET_I. In addition, as mentioned above, the sequence of visiting customers has already been determined by the static GVRP model, I need to design a backward function to deduce the$R_i\hat{}$ based on the feasible service start time at next customer.

Notation:

i and j = any pair of customers that would be visited in the given sequence
m and v = time interval number and vehicle speed
$R_j$is the current service start time at customer j and$R_j \le ET_j$and$R_j \le R_j\hat{}$.

$R_j \longleftarrow R_j\hat{}$

find time interval m, which meets $t^m \le R_j \le t_m$
$B_i \longleftarrow R_j - D_{ij}/v_m$

D $\longleftarrow D_{ij}$, t $\longleftarrow R_j$

While $B_i < t^m$ do
D ← D-(t-t^m)* v_m
t ← t^m
B_i ← t-D/v_{m-1}
m ← m+1
End While
R_i^ ← min(B_i-S_i,LT_I)
Output: R_i^

The next step is to define the whole optimization procedure as follows:

For each customer in the given delivery route, the earliest feasible service start time is determined by the beginning of the time window feasible service time could be obtained by running the above backward function.

Then, calculate each P_{ij}, when $B_i = t^m$. Calculate each when
Output the minimized P_{ij} and corresponding B_i.

Assuming the service time at each customer point is 5 minutes and the opening hours of the depot is from 9AM to 12PM, the optimization result is showed by the right table. Compared with the optimization result of the static GVRP, this model reduces 22g carbon emissions, occupying 0.9% of the total amount of carbon emissions(see table 8).

Table8 Optimized Departure Time and Waiting Time at Each Customer Point

	Sequence	Arrive Time	Wait Time	Departure Time
Route 1	2	9:00:42 AM	14.3	9:20:00 AM
	3	9:00:36 AM	0.0	9:05:36 AM
	4	9:07:06 AM	4.9	9:17:00 AM
	5	9:17:08 AM	47.9	10:10:00 AM
	6	10:10:26 AM	9.6	10:25:00 AM
	7	10:25:34 AM	0.0	10:30:34 AM
	8	10:31:01 AM	0.0	10:36:01 AM
	9	10:36:09 AM	1.2	10:42:21 AM
	10	10:47:03 AM	0.0	10:52:03 AM
	11	10:52:36 AM	0.0	10:57:36 AM
	12	10:57:41 AM	0.0	11:02:41 AM
	13	11:01:39 AM	0.0	11:06:39 AM
	14	11:07:58 AM	4.5	11:17:28 AM
	15	11:18:28 AM	7.1	11:30:34 AM
	16	11:32:22 AM	0.0	11:37:22 AM
	17	11:37:55 AM	0.0	11:42:55 AM
	18	11:42:56 AM	0.0	11:47:56 AM
	19	11:48:35 AM	0.0	11:53:35 AM
	Depot	11:55:00 AM		
Route 2	2	9:01:13 AM	19.8	9:25:00 AM
	3	9:25:12 AM	29.8	10:00:00 AM
	4	10:00:09 AM	34.8	10:40:00 AM
	5	10:41:23 AM	0.0	10:46:23 AM
	6	10:47:51 AM	0.0	10:52:52 AM
	7	10:52:57 AM	2.0	11:00:00 AM
	8	11:00:42 AM	0.0	11:05:42 AM
	9	11:06:07 AM	0.0	11:11:07 AM
	10	11:13:01 AM	2.0	11:20:00 AM
	Depot	11:21:03 AM		
Route 3	2	9:00:38 AM	25.2	9:30:50 AM
	3	9:31:07 AM	33.8	10:09:15 AM
	4	10:09:18 AM	0.7	10:15:00 AM
	5	10:15:14 AM	0.0	10:20:14 AM
	6	10:21:37 AM	0.0	10:26:37 AM
	7	10:27:17 AM	2.7	10:40:00 AM
	8	10:41:01 AM	0.0	10:46:01 AM
	9	10:46:20 AM	3.6	10:55:00 AM
	10	10:55:40 AM	0.0	11:00:40 AM
	11	11:01:09 AM	5.0	11:10:00 AM
	12	11:10:31 AM	0.0	11:15:31 AM
	Depot	11:16:02 AM		

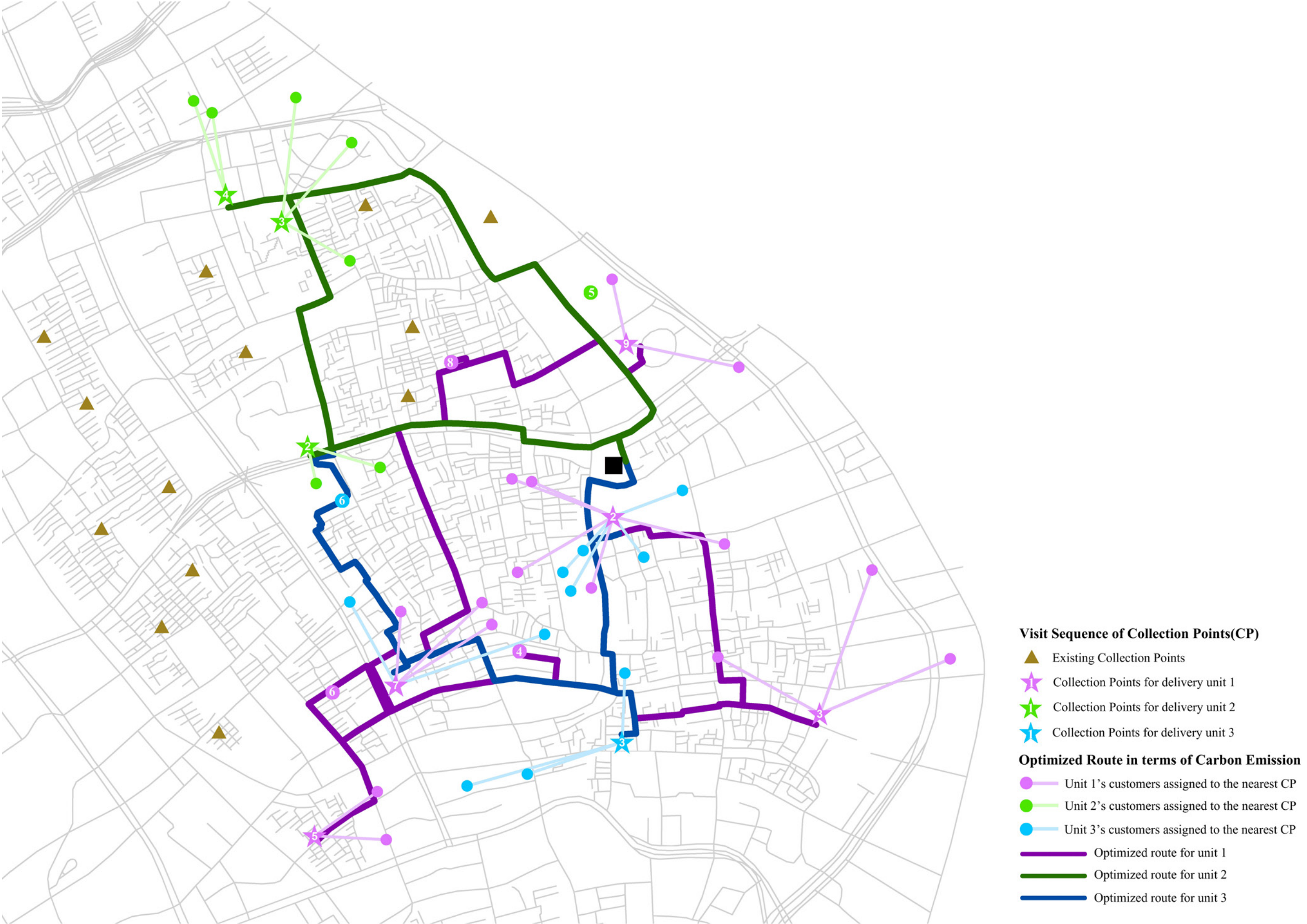
Advantages of Collection Points in Delivery System

Though after two tiers’ optimization (static model and dynamic model), the amount of carbon emissions during delivery process has been obviously reduced, the “door-to-door” (END) method is still not the most environment-friendly way. In fact, previous studies about delivery of books purchased online have justified the advantages of using collection points for picking up goods in reducing CO2 emissions. Kim and colleagues(2008, 2009) designed several collection points for picking up books by customers themselves in the United States. They compared the emissions of the proposed “sustainable networked delivery”(SND) system and “e-commerce networked delivery”(END) system in delivering 100 books to 100 customers. Their 2008 study indicated that the CO2 emissions of END are 12 times those of the SND system and their 2009 study showed that difference is 7 times. Zhang(2013) conducted a similar study in China, his study showed that though SND system in China has greater environmental impacts than that in UK and US, SND is much more environment-friendly than END system. In fact, building collection points in a highly densely-populated area like Shanghai is an ideal substitute of “door-to-door” delivery.

As mentioned in previous sections, there are some formal collection points built by express companies in the study area now. However, these points are exclusively used for collecting parcels within a particular area such as an office building or a campus. Actually, the initial reason for building such points is that couriers do not have permissions to enter these areas while the demand of these areas for parcels is high. However, in many cases, there are a lot of parcels need to be delivered to places close to collection points as well, but couriers cannot utilize those points. This section assumes a scenario that within an approachable buffer, all parcels are delivered to collection points(CP) instead of to customers one by one and customers walk to CP to pick up parcels by themselves. To simplify the scenario, following assumptions are made:

- All existing formal CPs are potentially selected for collecting parcels and all of them have the same weights. This means two things: no new CPs will be proposed; land use around CPs as well as distance between CP and customers will not affect customers’ willingness to access CPs.
- The rule of assigning parcels is 10-minutes-walking distance and parcels within this distance are assigned to the nearest CP. Walking distance is determined by the real street network rather than the planar distance on the map. In addition, all streets are assumed to have sidewalks.
- The capacities of CPs are assumed to be infinite. Actually, the result of this sections is expected to encourage express companies to expand existing CPs, so there is no need to limit capacities.
- Assumed customers and their information are the same with that used in GVRP model building, which means the inputs are still the 37 customers around depot 5.
- Delivery districts and delivery units still exist and the processes of assigning parcels in one unit to the nearest CP are separate. One CP could be shared by different units, but it have to be visited by different delivery vans.

Graph9　Advantages of Collection Points in Reducing Carbon Emissions



Graph 8 illustrates the optimization result of this section. Among the 18 customers of delivery unit 1, 15 customers were assigned to 5 collection points. Among the 9 customers of delivery unit 2, 8 customers were assigned to 3 collection points. Among the 11 customers of delivery unit3, 10 customers were assigned to 3 collection points. In total, 9 out of 22 existing formal collection points were selected to collect parcels in the simulation. Compared with the “door-to-door” model, collection points system reduces 332 g carbon emissions, occupying 19.7% of the total amount. In addition, this model also saves 19% travel time. This result demonstrates that END system has much better performance in reducing carbon emissions than either static GVRP model or dynamic GVRP model. Though this project does not research the exact relationship between collection points’ effectiveness and population density, Shanghai, or most Chinese cities, is still the best place to apply collection points delivery system.

Table9　Comparision between “Door-to-Door” Delivery and “Collection Points” Sys-

		Virtual Distance(meters)	Carbon Emissions(g)	Travel Time(minutes)
Route 1	“Door-to-Door”	12,479	774	14.9
	Collection Points	10,661	661	12.8
Route 2	“Door-to-Door”	7,143	443	8.6
	Collection Points	5,357	332	6.4
Route 3	“Door-to-Door”	7,526	467	9.0
	Collection Points	5,792	359	6.9

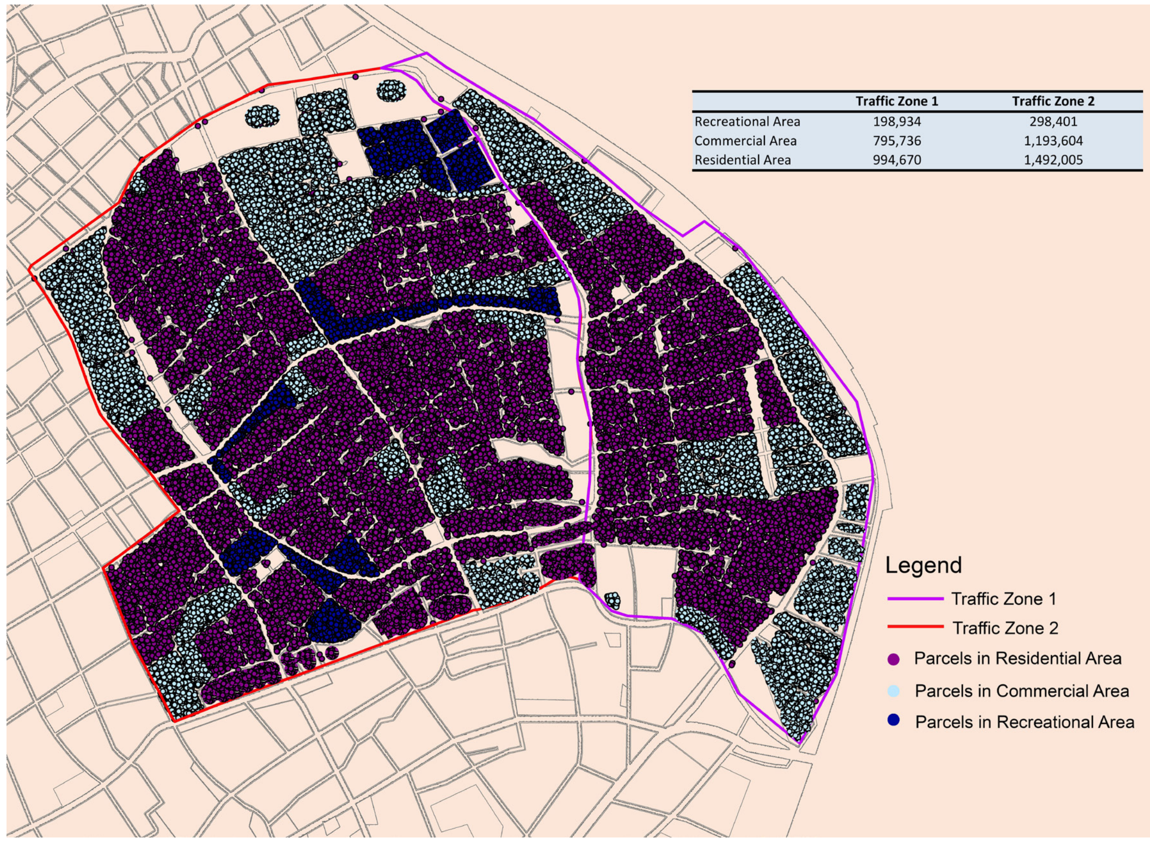
Best location of Collection Points

The major problem to apply collection points in the study area is to figure out the best location of proposed collection points. In fact, even though express companies are willing to eliminate the service limitations of their formal collection points, the location of existing formal CPs may not be the most suitable. Like illustrated in previous sections, there are plenty of informal collection points as well. Assuming the number of formal collection points would not change, but some formal CPs could be replaced by informal CPs, would both express companies and the society be better off? On the one hand, express companies do not need to provide more money to for additional CPs, so they are willing to apply such exchange. On the other hand, some informal CPs locate in more densely-parcel-concentrated areas, so they could serve much more customers and subsequently reduce carbon emissions. This section proposes a method for express companies to best locate CPs based on the distribution of parcels through a whole year, which is specifically described in following procedure:

Step 1. Generation of Customers

Since this section needs a big sample of parcels, we could not use the assumed customers in previous sections. Instead, we need to generate parcels by land use. According to the annual logistic report of Shanghai, the parcels’ distribution in 2013 could be summarized as 50% in residential area, 40% in commercial area and 10% in recreational area. In addition, when simulate the parcels, we should tease out the residential areas filled with old people since these areas’ demand for parcels is too tiny to be calculated. Since the study area has population of 198,934 and each person produces 25 parcels/year, the parcels’ distribution by land use in the two involved traffic zones is listed by graph 10:

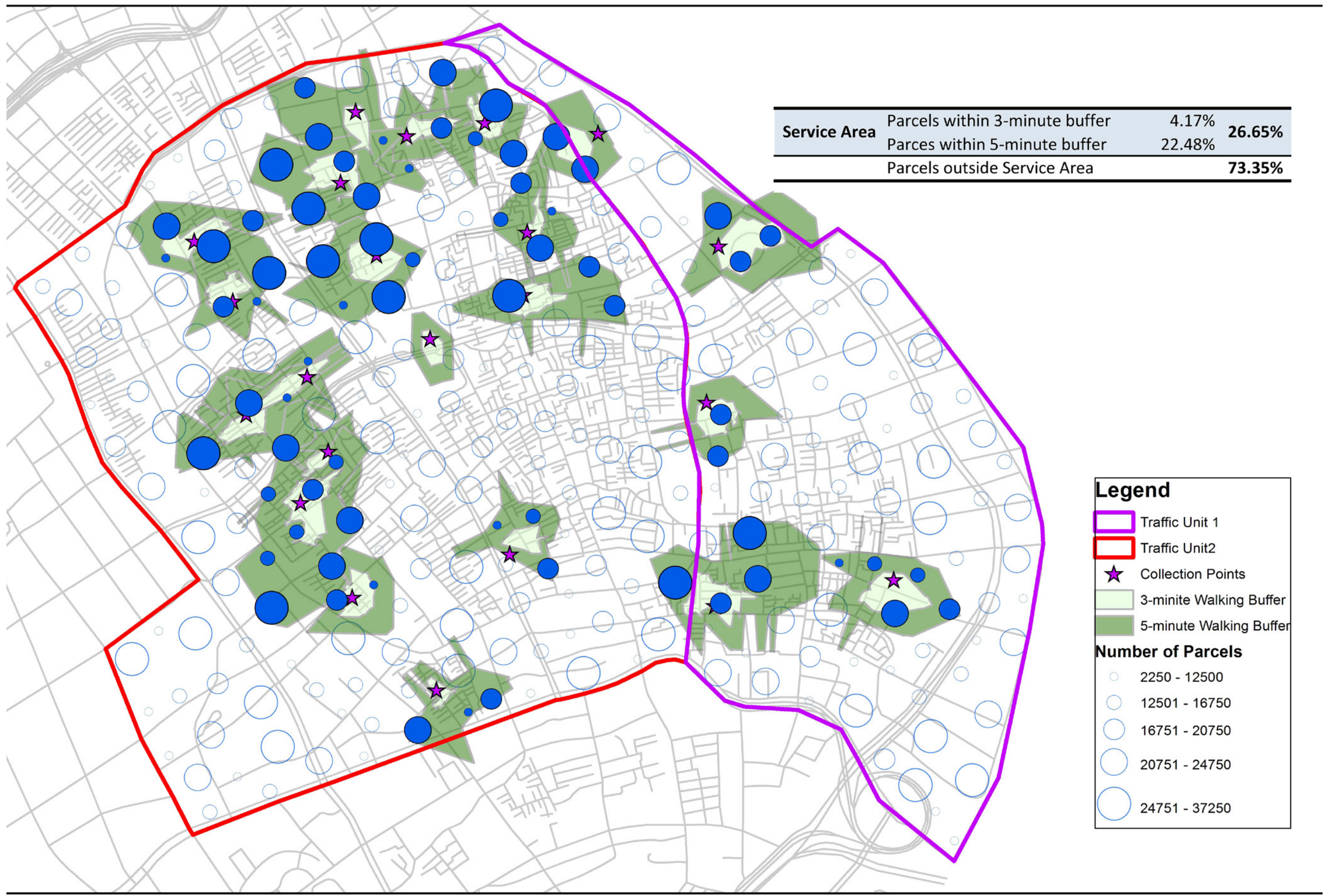
Graph10　Distribution of E-commerce Customers



Step 2. Evaluation of Existing Formal Collection Points' Effectiveness

To simplify the analysis, parcels within 50 meters were integrated as a centroid and different centroids have different counts of parcels. As demonstrated by the following map, service areas of each existing collection point are defined by 3-minute walking and 5-minute walking buffer. Solid blue circles denote parcel centroids that fall within service areas while hallow circles denote parcel centroids outside service areas.(see graph 11) Currently, 26.65% of total parcels per year can be served by existing CPs.

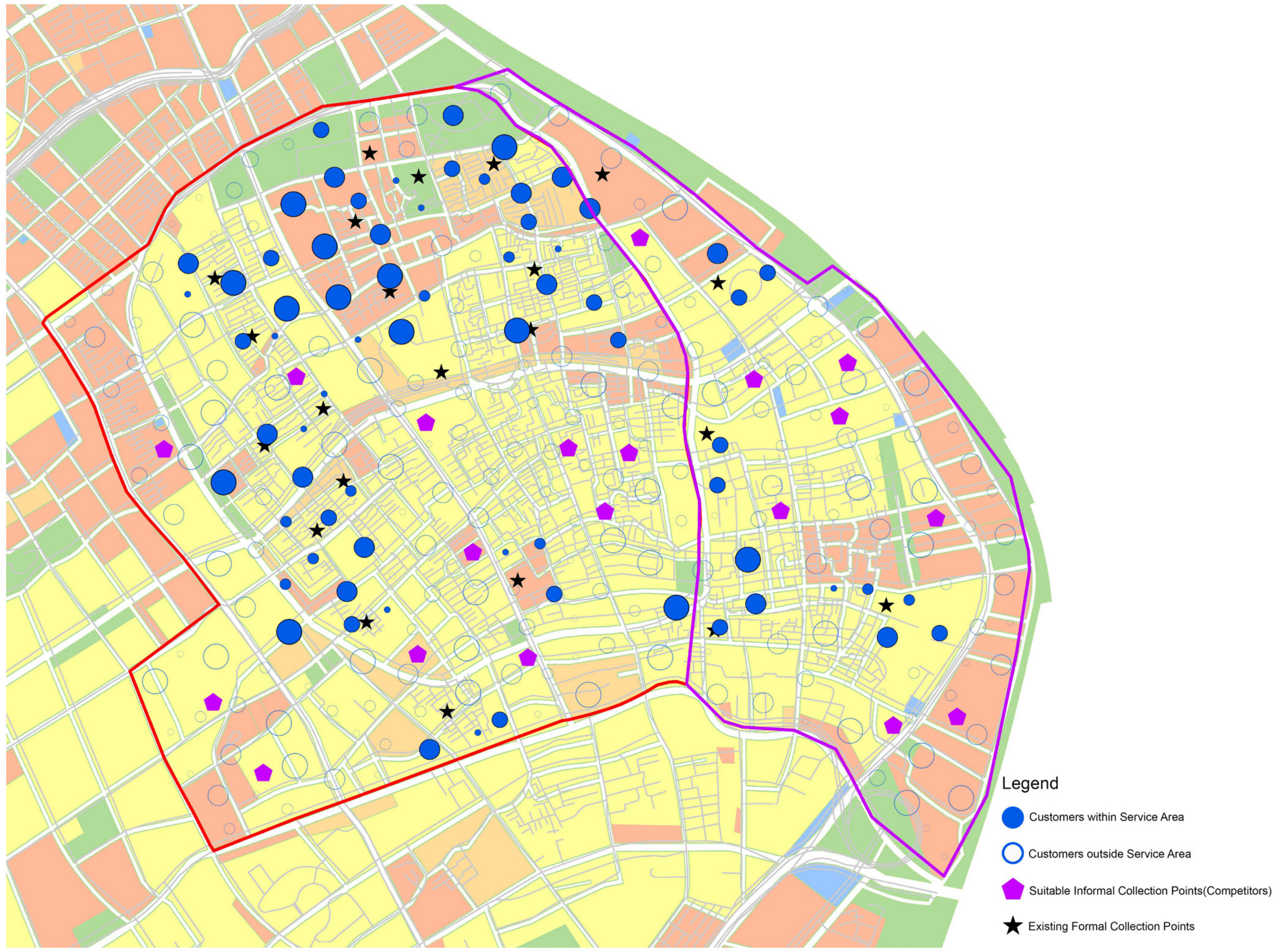
Graph11 Evaluation of Existing Collection Points' Effectiveness



Step 3. Competition between Formal and Informal Collection Points

Competitors for proposed collection points are selected from informal collection points based on their willingness to build cooperative relationships with express companies. In fact, my cooperation planning company CAUPD contacted each existing informal collection point that locate within areas where customers outside the service area of formal collection points concentrate. In other words, alternative collection points represent those with willingness to collect parcels from broader areas rather than their own communities as long as express companies provide subsidize were listed as competitors.(see graph 12)

Graph12 Distribution of Alternative Collection Points



Finally, a 5-minute-walking service area was created for both existing CP and competitor. The counts of customers that fall within each service area was calculated, and existing CPs as well as competitors were ranked based on their counts of customers from the largest to the smallest.(see Graph 130 The top 22 ones were selected as proposed collection points. As illustrated by the following map, 10 out of 22 existing collection points were remained and 12 competitors were selected. That means the express company should replace the knockout formal CPs with selected informal CPs. After this exchange, the total rate of customers that fall within 5-minute-walking service area increases from 26.65% to 37.5%. In fact, the evaluation of CPs and relocation of CPs could be updated every year to respond to the changing logistic market.

Conlusion and Future Work

This research work has shed light on three aspects of "Last Mile" delivery:

First, the study of current parcel-delivery system in Shanghai has provided a good understanding of how customers receive parcels in the last-mile part. For example, "door-to-door" delivery in Shanghai includes a soft time window which is exclusively set by couriers, but the customers' "not-at-home" frequency is quite low. This finding contradicts with our traditional wisdom. Another important finding is all existing collection points were built by courier services and the construction reason is that couriers do not have permissions to enter particular districts. This finding can explain why the effectiveness of collection points in Shanghai is limited: first, though collection points have abundant storage space, existing CPs exclusively serve customers within particular area rather than a broader area; second, the function of existing CPs is single and CPs could have more diverse activities. In fact, it is prevalent for courier services to cooperate with gas stations or grocery stores.

Second, this project proposed an optimization model to optimize sequence of customers being visited, delivery routes and departure time at each customer point in terms of minimizing carbon emissions. Specifically, the static GVRP model could reduce 3% carbon emissions generated by the currently prevalent "the shortest distance" delivery strategy. Besides, the dynamic GVRP model could further reduce 0.9% carbon emissions. Courier services could use the proposed optimization model to plan delivery routes before every day's delivery tasks and print them out for couriers.

Third, this project justified the importance of promoting collection point system and presented a practical framework for locating new CPs. On the basis of the proposed optimization model, CP system could further 19% carbon emissions and 21% travel time. In fact, given the compact environment and concentration of E-commerce customer, Shanghai may be the most ideal place to apply collection point system. Currently, 26.65% of customers fall within 5-minute walking buffer of CPs which was regarded as a reasonable service area of the CP. To maximize the coverage of CPs' service areas, 10 out of 22 existing collection points need to be replaced by alternatives. All the alternatives which expressed the willingness to cooperate with courier services are grocery stores. The establishment of cooperative relationship between CPs and courier services is expected to benefit both of them and the whole society.

Due to the limited data source and time, the project has several limitations and some future work has to be done. For the VRP optimization model part, to more precisely quantify effects of multiple factors on carbon emissions, a large-sample survey needs to be conducted by the transportation department of Shanghai. In addition, courier services should provide more data, including number of employers and types of delivery vans. These data are useful to differentiate carbon emissions of different scenarios. For the location-allocation part, the local government can collect data about age and gender composition of E-commerce customers, population by small tracts, and sidewalk conditions near CPs. All of these data play a significant role in analyzing the willingness of customers to get access CPs. In one word, this project created a basic framework to incorporate environmental costs into the "Last Mile" optimization, the establishment of a really "Green Last Mile" delivery system in Shanghai needs cooperation among multiple groups.

Bibliography

1. Jiang, Xiaomei. "Development of China's Logistics Market." Current Chinese Economic Report Series Contemporary Logistics in China (2014): 21-41. Web. <http://www.atkearney.com/documents/10192/5274810/Chinas+E-commerce+market-+The+Logistics+Challenges.pdf/76119ca8-ce91-464e-b8da-b9ad301586a3>.

2. Gevaers, Roel, Eddy Van De Voorde, and Thierry Vanelslander. "Characteristics and Typology of Last-mile Logistics from an Innovation Perspective in an Urban Context." City Distribution and Urban Freight Transport (2011): n. pag.

3. Kontoravdis, George, and Jonathan F. Bard. "A GRASP for the Vehicle Routing Problem with Time Windows." ORSA Journal on Computing 7.1 (1995): 10-23. Web.

4. Figliozzi, Miguel Andres. "An Iterative Route Construction and Improvement Algorithm for the Vehicle Routing Problem with Soft Time Windows." Transportation Research Part C: Emerging Technologies 18.5 (2010): 668-79. Web.

5. Figliozzi, Miguel Andres. "An Iterative Route Construction and Improvement Algorithm for the Vehicle Routing Problem with Soft Time Windows." Transportation Research Part C: Emerging Technologies 18.5 (2010): 668-79. Web.

6. Ichoua, Soumia, Michel Gendreau, and Jean-Yves Potvin. "Vehicle Dispatching with Time-dependent Travel times." European Journal of Operational Research 144.2 (2003): 379-96. Web.

7. Liu, Wan-Yu, Chun-Cheng Lin, Ching-Ren Chiu, You-Song Tsao, and Qunwei Wang. "Minimizing the Carbon Footprint for the Time-Dependent Heterogeneous-Fleet Vehicle Routing Problem with Alternative Paths." Sustainability 6.7 (2014): 4658-684. Web.

8. Zhang, Liyi, Ying Wang, Teng Fei, and Hongwei Ren. "The Research on Low Carbon Logistics Routing Optimization Based on DNA-Ant Colony Algorithm." Discrete Dynamics in Nature and Society 2014 (2014): 1-13. Web.

9. "Supply Chain Performance Evaluation and Design." Stochastic Discrete Event Systems (2008): 307-23. Web.

10. Huang, Yijun, and Kaikai Yin. "Evaluation Indexes of the Third-Party Logistics Providers' Logistics Capability in B2C E-commerce Mode." Liss 2013 (2013): 181-89. Web.

11. Yu, Changhui. "Research and Design of Logistics Management System Based on Internet of Things." 2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC) (2011): n. pag. Web.